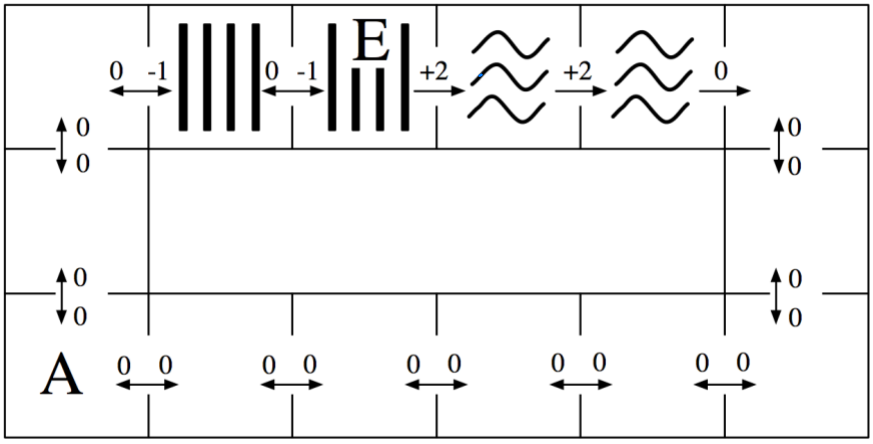
CSE 415 Computer Science and Engineering, University of Washington

Name: Travis Xie

# Waterpark World
## Part I: MDP Values

In the "Waterpark World" environment, an agent can move either clockwise or counterclockwise, except in 3 states where the agent must go clockwise due to the presence of a water slide. There are two states that represent being on a ladder leading up to the slide. Rewards for moving are 0, except that going down either portion of the slide gives a +2 reward (fun!) and climbing either part of the ladder has a reward of -1 (feels like work!). The agent has to move on every turn. Assume noise = 0.0 (i.e., actions are deterministic).




(a) How many distinct policies are possible for this MDP? $2^{11}$

(b) Fill in the blank cells of this table with values that are correct for the corresponding function, discount, and state. Hint: You should not need to do substantial calculation here.

| | $\gamma$ | $s = A$ | $s = E$ |
|---|---|---|---|
| $V_3(s)$ | 1.0 | 0 | 4 |
| $V_{10}(s)$ | 1.0 | 2 | 4 |
| $V_{10}(s)$ | 0.1 | 0 | 2.2 |
| $Q_1(s, \text{left})$ | 1.0 | 0 | 0 |
| $Q_{10}(s, \text{left})$ | 1.0 | 2 | 3 |
| $V^*(s)$ | 1.0 | $\infty$ | $\infty$ |
| $V^*(s)$ | 0.1 | 0 | 2.2 |

(V-ST-19-11-01. This exercise is based on an example used at U.C. Berkeley)